PROCEEDINGS OF TDWG

Biodiversity Information Science and Standarts 1: e20483
doi: 10.3897/tdwgproceedings.1.20483

OPEN ACCESS

Conference Abstract

Environmental samples, eDNA and HTS libraries – data standard proposals from the Global Genome Biodiversity Network (GGBN)

Gabriele Droege[‡], Jonas Zimmermann[‡], Tim Fulcher[§], Sietse Van der Linde[|], Walter G Berendsohn[‡]

‡ Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin, Berlin, Germany
§ Royal Botanic Gardens Kew, London, United Kingdom
| Forest Research – Forestry Commission, Farnham, United Kingdom

Corresponding author: Gabriele Droege (g.droege@bgbm.org)

Received: 21 Aug 2017 | Published: 21 Aug 2017

Citation: Droege G, Zimmermann J, Fulcher T, Van der Linde S, Berendsohn W (2017) Environmental samples, eDNA and HTS libraries – data standard proposals from the Global Genome Biodiversity Network (GGBN). Proceedings of TDWG 1: e20483. https://doi.org/10.3897/tdwgproceedings.1.20483

Abstract

The GGBN Data Portal (http://www.ggbn.org, Droege et al. 2014) has established standardised data flows for genomic DNA samples, including voucher specimens, tissue samples, DNA samples as well as resulting sequences and publications. Dealing with different types of DNA (aDNA, gDNA, eDNA) is essential and closely related to user-friendly search and display functionalities. GGBN aims both at preserving voucher specimens of all kinds of DNA as well as making these important data accessible on the Internet.

In addition to genomic DNA, the development and use of high-throughput-/next-generation-sequencing (HTS formerly designated NGS) has outstripped current plans of SYNTHESYS and GGBN to join natural history collection data with DNA and tissue collection data. HTS libraries can be considered as a preparation of the genetic material of a single organism or of multiple organism (e.g. from an environmental mixed sample). From that point of view, they are the actual physical molecular representation of a specimen or sample. However, these libraries come with specific adaptors that limit their transferability to other sequencing systems. The libraries are prepared at great expense, but frequently are only used for a

© Droege G et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

single project, not making use of additional useful information that could potentially be generated. To increase the potential of the HTS libraries to be used for multiple projects they have to be discoverable via published metadata. Optimally, HTS library metadata will include specific standardized keywords (by e.g. organism, HTS method etc.).

Here we present our ideas and a prototype for eDNA samples and HTS libraries based on the GGBN Data Standard (Droege et al. 2016). A use case collection from animals, diatoms, fungi and plants has been developed and is available in the GGBN Sandbox (see more details at http://wiki.ggbn.org/ggbn/Use_Case_Collection). These examples will be improved upon and kept on-line at least until 2020, so progress can be observed.

Keywords

GGBN Data Standard, eDNA, environmental samples, HTS/NGS libraries

Presenting author

Walter G Berendsohn

Funding program

The authors thank the European Commission, the German Research Foundation (DFG), and the National Museum of Natural History (Smithsonian Institution) Global Genome Initiative for funding this work within SYNTHESYS III (312253), DFG-GGBN (GU 1109/5-1, GE 1242/13-1), DFG-DNA-Bank-Netzwerk (INST 1039/1-1, INST 17818/1-1, INST 427/1-1, INST 599/1-1), DFG-GFBio (GU 1109/3-1) and DFG-BiNHum (BE 2283/8-1).

References

- Droege G, Barker K, Astrin J, Bartels P, Butler C, Cantrill D, Coddington J, Forest F, Gemeinholzer B, Hobern D, Mackenzie-Dodds J, Tuama ÉÓ, Petersen G, Sanjur O, Schindel D, Seberg O (2014) The Global Genome Biodiversity Network (GGBN) Data Portal. Nucleic Acids Research 42: D607-D612. https://doi.org/10.1093/nar/gkt928
- Droege G, Barker K, Seberg O, Coddington J, Benson E, Berendsohn WG, Bunk B, Butler C, Cawsey EM, Deck J, Döring M, Flemons P, Gemeinholzer B, Güntsch A, Hollowell T, Kelbert P, Kostadinov I, Kottmann R, Lawlor RT, Lyal C, Mackenzie-Dodds J, Meyer C, Mulcahy D, Nussbeck SY, O'Tuama É, Orrell T, Petersen G, Robertson T, Söhngen C, Whitacre J, Wieczorek J, Yilmaz P, Zetzsche H, Zhang Y, Zhou X (2016) The Global Genome Biodiversity Network (GGBN) Data Standard specification. Database 2016: baw125. https://doi.org/10.1093/database/baw125